ISSN No. 0976-5697

Volume 5, No. 3, March-April 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

The Recognition of Holy Quran Reading types "Rewaih"

Talaat Wahbi
Department of Computer Science,
Sudan University of Science and Technology, Sudan

Ashwag Gadeed*
Department of Software Engineering,
Sudan University of Science and Technology, Sudan

Abstract: This paper investigates recognition of a new trend in the holly Quran research it is Rewaih recognition. A data set of 78 verses (Aiah) was collected for a one reciter for two reading types Hfs Rewaih and Dori Rewaih. The system was trained using 38 verses, a separate set of 40 verses has been used for testing. For each Rewaih an HMM model has been trained. The test recognition rates for both Rewaih were 100%.

Keywords: The Holy Quran; Automatic speech recognition; Cepstral feature extraction; mixtures of Gaussian; Hidden Markov model .

I. INTRODUCTION

Automatic speech recognition has been an active research topic for more than four decades. With the advent of digital computing and signal processing, the problem of speech recognition was clearly posed and thoroughly studied. These developments were complemented with an increased awareness of the advantages of conversational systems. The range of the possible applications is wide and includes: voice-controlled appliances, fully featured speech-to-text software, automation of operator-assisted services [1], and voice recognition aids for the handicapped, etc.

Automatic speech recognition (ASR) is having several applications that use this technology. One of the important Arabic applications is recognition based on holy Quran. There are many difficulties begin when dealing with the specialties of the Arabic language in Holy Quran, due to the differences between written and recite the Al-Quran

The holy Quran is revealed in Arabic for that it is the official language of many countries and has many different, geographically distributed spoken varieties, some of which are mutually unintelligible. Modern Standard Arabic (MSA) has basically 34 phonemes, of which six are vowels, and 28 are consonants. A phoneme is the smallest element of speech units that makes a difference in the meaning of a word, or a sentence. The correspondence between writing and pronunciation in MSA falls somewhere between that of languages such as Spanish and Finnish, which have an almost one-to-one mapping between letters and sounds ,and languages such as English and French, which exhibit a more complex letter-to-sound mapping [2].

So the holy Quran text reading is not same as Arabic known today. Some rules are implemented to ensure the correct reading of the Holy book. These rules have come to be known as Tajweed rules [3] [4]. The recitation of Holy Quran by use of Tajweed rules is an art. The recites' follow the Tajweed rules to build their recitation attractive [3]. If someone wants to learn the Tajweed then he/she has to consult an expert Hafiz at some learning institute. Learning is done manually as the student and recite have to sit face to face and the learner recites and the expert's points out and corrects the mistakes if any occur. Presently, there exist institutes and websites that teach the recitation in manual method [5]. Tajweed can be differentiated depending on reading holy Quran. It can be read in 10 different ways named 'Rewaih', of course some Rewaihs are different from others.

The main problem here is the accessibility and feasibility of reading type "Rewaih" expert (Hafiz). Availability of Hafiz in countries where Islam is not a dominant religion can be quite less to come across. For that this research done to solve the reading type problem [5].

Researches in the field of the holy Quran have not yet become mature independent applications until now, But many serious contributions in [6] [7] [8] [9] [10]. Many problems face these types of research. This is mainly because of neglecting of the Holy Quran recitation methods.

The following points will clear the understanding of Holy Quran reading:

a. The holy Quran reading methodology From "Tajweed" definition which is the science of knowing Arabic letters' rights and dues from place of pronunciation and their features e.g. "Baqdaian" rule (El Qaida El baqdaiah) has many steps for reading: Firstly, need pronunciation of Arabic letters correctly from the right place (Makhraj). Secondly, Learning Arabic letters' features (ElSefat). Thirdly learning Tajweed rules (Qnaa, Ezhar, etc…). Fourthly, Correcting recitation (Eltasheeh and Eldabt).
b. Apply Acoustic models and pattern recognition techniques' to detect special problem types for special type e.g. Problem of " Tajweed El Noon El Saknha and El Tanween" for beginners (beginners mean who learn Arabic letters places and feature) and not for all Holy Quran readers. Also the "Qalqal" Rules in Start and middle letters for recites who know letters place and features, etc….
c. The data set: This type of researches needs many different data sets like: famous recites, known rules of each "Rewaih", common reading mistakes and special reading types problem (e.g. Loud Sounds) data sets.

Some good datasets founded like [11], but it is for ten reciters and one chapter and one reading type. Really, there is a short in holy Quran prepared dataset for recognition research, especially speech data.

The most dominant approach for SR system is the statistical approach Hidden Markov Model (HMM), trained on corpora that contain speech resource from a large number of speakers to achieve acceptable performance.

The rest of the paper is organized as follows: Section 2 briefly introduces HMMs. Section 3 describes in general the dataset. Section 4 illustrates and outlines the results achieved by the experiments performed. Finally, a conclusion is drawn with future work outlooks in section 5.

II. HIDDEN MARKOV MODELS

Consider a discrete time Markov chain with a finite set of states' $S = \{s_1, s_2, ..., s_N\}$. An HMM is defined by the following compact notation to indicate the complete parameter set of the model $\lambda = (\Pi, A, B)$ where Π, A and B are the initial state distribution vector, matrix of state transition probabilities and the set of the observation probability distribution in each state, respectively.

$\Pi = [\pi_1, \pi_2, ..., \pi_N]$,

$\pi_i = P(q_1 = s_i)$,

$A = [a_{ij}]$,

$a_{ij} = P(q_{t+1} = s_j \mid q_t = s_i)$, $1 \le i, j \le N$, $s_i, s_j \in S$, $t \in \{1, 2, ..., T\}$.

The observation at time t, O_t, may be a discrete symbol (Discrete HMMs (DHMMs) case), $O_t = v_k$, $v_k \in V = \{v_1, v_2, ..., v_M\}$, or continuous, $O_t \in R^K$. For a discrete observation, v_k will be written as z for simplicity.

The observation matrix B is defined by $B = [b_i(O_t)]$, where $b_i(O_t)$ is the state conditional probability of the observation O_t defined by $b_i(O_t) = P(O_t = z \mid q_t = i)$, $1 \le i \le N$, $1 \le z \le M$. For a continuous observation (Continuous HMMs (CHMMs) case), $b_i(O_t)$ is defined by a finite mixture of any log-concave or elliptically symmetric probability density function (pdf), e.g. Gaussian pdf , with state conditional observation mean vector μ_i and state conditional observation covariance matrix Σ_i, so B may be defined as $B = \{\mu_i, \Sigma_i\}$, i = 1, 2, ..., N. The model parameters constraints for $1 \le i, j \le N$ are

$$\sum_{i=1}^{N} \pi_i = 1, \sum_{j=1}^{N} a_{ij} = 1, a_{ij} \ge 0, \sum_{k=1}^{M} b_i(O_t = z) = 1 \; or \; \int_{-\infty}^{\infty} b_i(O_t)\, dO_t = 1 \qquad (1)$$

In the case of left-to-right HMMs, multiple observation sequences must be used. Unlike the ergodic HMM case, the left-to-right HMM necessitates taking into account that the initial state probability is always equal to one. Thus, all the sequences are supposed to start with the same state. In this work, the specific structures are taken into account by weighting the values of the elements of the matrix A. The weights are defined by zeros and ones elements corresponding to the structure of the matrix A. In general, at each instant of time t, the model is in one of the states i, $1 \le i \le N$. It outputs O_t according to a discrete probability (in the DHMMs case) or according to a continuous density function (in the CHMM case) $b_j(O_t)$ and then jumps to state j, $1 \le j \le N$ with probability a_{ij}. The state transition matrix defines the structure of the HMM. The model λ may be obtained off-line using some training algorithm. In practice, given the observation sequence $O = \{O_1 O_2 ... O_T\}$, and a model λ, the HMMs need three fundamental problems to be solved. Problem1 (Evaluation problem) is how to calculate the likelihood $P(O|\lambda)$. The solution to this problem provides a score of how O belongs to the model λ. Problem2 (Decoding problem) is how to determine the most likely state sequence that corresponds to O.

The solution to this problem provides the sequence of the hidden states corresponding to the given observation sequence O. Problem3 (Learning problem) is how to adjust the model λ in order to maximize $P(O|\lambda)$. This is the problem of estimating the model parameters given a corpus of training observations sequences. Problems 1 and 2 are solved in the decoding stage using either the forward probability or the Viterbi decoding algorithm, while problem 3 is solved during the training stage using either a conventional algorithm such as the Baum-Welch algorithm and the Viterbi-based algorithm, known as segmental K-mean in the case of CHMMs or some other new training algorithm[12] [13].

The three problems will be look like the following for an image recognition example as follow:

a. The evaluation and decoding problems: Calculate the probability of specific patterns' frames (the first five frames V^5 set) when entered to states set (w_1, w_2, w_3, w_4) in time from 0 ~ 4. After calculation of probability get the best path of states that can get the specific patterns' frames (max probability) to solve decoding problem. Figure 1 show demonstration of those processes.

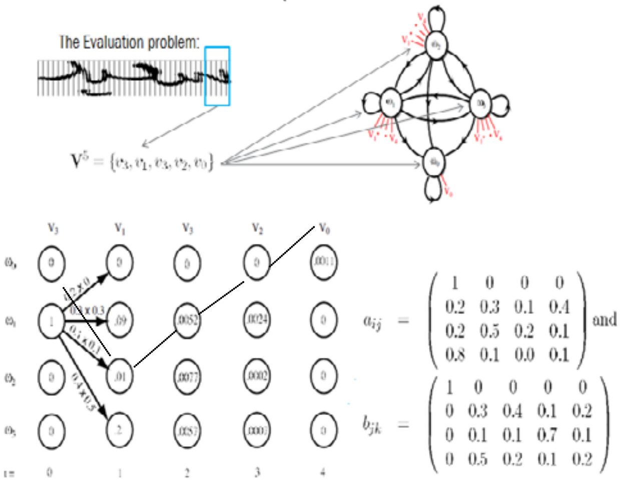

Figure 1. Evaluation problem: Probability calculation for matrix a and b. decoding problem: max probability selection shown by dotted line .

III. CONTRIBUTION

In this paper, we propose an efficient scheme based on HMM, aiming at recognizing the reading types (Rewaih) mainly Hfs and Eldori. This scheme not only differentiates Hfs from Eldori but also recognizes the verses that have typical readings in both Rewaih.

IV. DATASET

Two Reading types were used, the first one is Hfs from Asim for Shikh Moftah AlSultany Musshaf, and the second is Eldori from Elksaiee also for same Shikh. Three soras were selected from chapter (Goze) 30 as shown in table 1 for both reading.

Table 1. Selected Soras

Sort of sora in holy Quran	Sora name	Number of verses
91	Elshams	15
92	Ellail	21
111	Elmasad	5

Figure 2 shows Elshams and Elmasad soras and the differences between the two readings are shown as a colored letters.

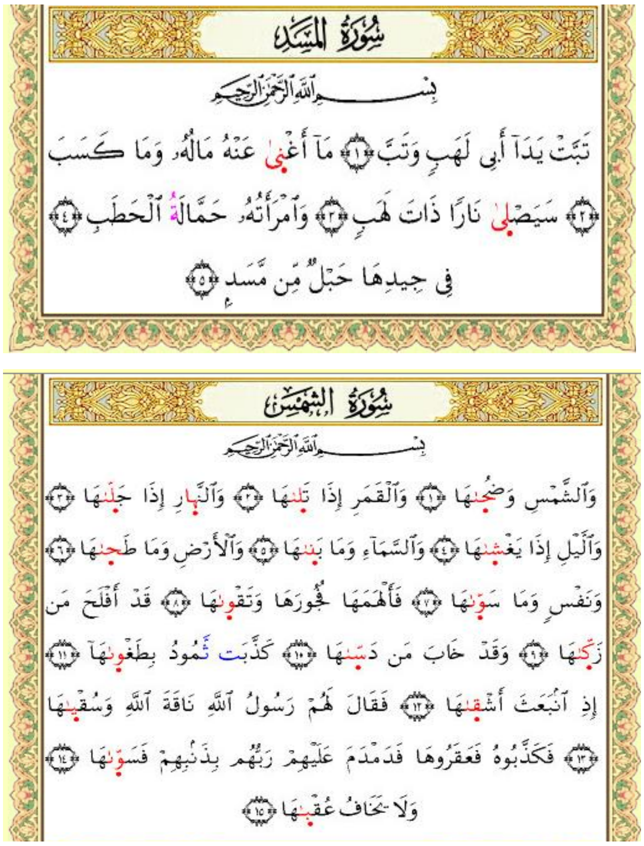

Figure 2. Shows the letter with “emalh” in Eldori in red color and other differences with magenta from (NOON Quran web site [14]).

Each sora is divided to its verses in each wav file after measuring active level for speech in ITU-T P.56 [15]. Table 2 shows train and test data from sora’s verses for both reading types.

Table 2. Test and train data

Sora name	**Number of verses**	**Train data**		**Test data**	
		Hfs	**Eldori**	**Hfs**	**Eldori**
Elshams	15	0	0	15	15
Ellail	21	16	16	5	5
ElMasad	5	3	3	0	0

V. EXPERIMENTS AND RESULTS

For recognition of holy Quran Rewaih using HMM, a recognition system that used Mel-Frequency Cepstral Coefficient (MFCC) as a feature extraction method is constructed. MFCC is considered the best technique because behavior of acoustic system remains unchanged during transferring the frequency from linear to non-linear scale [5]. The data set is recorded in studio so the noise level was very low and this calculated from active level detection, for that many techniques to avoid noise are not used. The main recognition system stages are described below:

A. *Stage one Frame blocking:*

Since the vocal tract moves mechanically slowly, speech can be assumed to be a random process with slowly varying properties [5]. Hence, the speech is divided into overlapping frames by length 1024 and half frame overlapping. The speech signal is assumed to be stationary over each frame and this property will prove useful in the following steps.

B. *Stage two Windowing:*

To minimize the discontinuity of a signal at the beginning and end of each frame, each frame fragmented to window to increase the correlation of MFCC mel cepstrum estimates between consecutive frames [15]. A Hamming window in time domain was used. Hamming window form is

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad 0 \le n \le N-1 \tag{2}$$

C. *Stage three Features extraction:*

The most prevalent and dominant method used to extract spectral features is calculating Mel-Frequency Cepstral Coefficients (MFCC). It is based on the frequency domain of Mel scale for human ear scale. Mel-scale has been used to perform filter bank processing to the power spectrum. It had been performed after windowing process and FFT had been implemented. Mel cepstrum with 12 coefficients for selected 256 sample frames using 32 filter banks with high and low filters with values 0.5 and 0 as a fraction of frequency [15].

D. *Stage four Classification:*

In the classification stage, 2 HMMs with 3 states for both reading types was used. Since no general parametric families of such distributions are known, in the continuous case, probability distributions are usually approximated via state-specific mixtures of Gaussians. HMM was modeled using 3 Gaussian Mixtures. Table 3 shows the results of train and test data.

Table 3. Train and test recognition rates for Hfs and Eldori readings

	Hfs	**Eldori**
Train	84.21%	100%
Test	100%	100%

The error in the train dataset testing is due to training the two HMMs with three verses that haven’t “Emalh” in the “Eldori” reading types. This means that there is another level of measuring the difference between readings types in case that “Emalh” is not found.

VI. CONCLUSION AND FUTURE WORK

In this paper the first steps towards developing Quranic Rewaih recognition using HMM was reported. A small corpus collected and used to train and test for an acoustic model. The result of test data for both Rewaih was 100%. Many future works can be done, as an example a work will specialize to recognize the differences like “Emalh” alone for acoustic model. Another issue in Quran research is that there is no standard open dataset for one Rewaih, So large and standard dataset including all reading types is the first step to go further in this area.

VII.REFERENCES

[1] D. Deore and S.Shrivastav, A Statistical approach in Natural Language Processing for Detecting Spoken numbers, International Journal of Research and Reviews in Signal Acquisition and Processing, Vol. 1, No. 1, March 2011.

[2] Nizar Y. Habash, "Introduction to arabic natural language processing" 2010 by Morgan & Claypool.

[3] H. Tabbal, W. El-Falou and B. Monla, Analysis and Implementation of a "Quranic" verses delimitation system in audio files using speech recognition techniques. Proc. of the IEEE Conf. of 2nd Information and Communication Technologies 2006; 2(1): 2979 – 2984.

[4] M.S. Bashir, S.F. Rasheed, M.M.Awais, S. Masud and S. Shamail, Simulation of Arabic Phoneme Identification through Spectrographic Analysis. CS, LUMS, Lahore, Pakistan,2003.

[5] Aslam Muhammad, Zia ul Qayyum, Waqar Mirza M. Saad Tanveer, Martinez-Enriquez A.M., Afraz Z. Syed: EHafiz: Intelligent System to Help Muslims in Recitation and Memorization of Quran. Life Science Journal. 2012;9(1):534-541] (ISSN:1097-8135).

[6] Noor Jamaliah Ibrahim, Zaidi Razak, Mohd Yakub, Zulkifli Mohd Yusoff, Mohd Yamani Idna Idris and Emran Mohd Tamil, Quranic verse Recitation feature extraction using Mel- Frequency Cepstral Coefficients (MFCC). Proc. the 4th IEEE Int. Colloquium on Signal Processing and its Application (CSPA) 2008, Kuala Lumpur, MALAYSIA. (2008).

[7] Program j-QAF sentiasa dipantau, Berita Harian, Online (2005).

[8] Zaidi Razak, Noor Jamaliah Ibrahim, Mohd Yamani Idna Idris, Emran Mohd Tamil, Mohd Yakub, Zulkifli Mohd Yusoff and Noor Naemah Abdul Rahman, Quranic Verse Recitation Recognition Module for Support in j-QAF Learning: A Review. IJCSNS International Journal of Computer Science and Network Security 2008; 8(8).

[9] الحاج، يحيى، منصور الغامدي، محمد الكنهل، عبد الله الأنصاري (1433هـ) نحو مصحح آلي للتلاوة القرآنية مدمج في بيئة حاسوبية للتعلم الذاتي للقرآن الكريم .المجلة العربية لعلوم وهندسة الحاسوب 11: 01--15.

[10] الحاج، يحيى، منصور الغامدي، محمد الكنهل، عبد الله الأنصاري (1431هـ) التعرف الآلي على الأصوات القرآنية أثناء التلاوة .الندوة الدولية السادسة لعلوم وهندسة الحاسوب .الحمامات، تونس. 6-7 جمادى الآخرة 1431هـ.

[11] الحاج، يحيى، منصور الغامدي، محمد الكنهل، عبد الله الأنصاري (1430هـ) ذخيرة صوتية لجزء من القرآن الكريم (النبأ). ندوة القرآن الكريم والتقنيات المعاصرة .المدينة المنورة. 24-26 شوال 1430هـ.

[12] Rabiner, L.R., "A tutorial on hidden Markov model and selected applications in speech recognition", IEEE Proc., 77, 1989, pages 257-386.

[13] Tarik Al-ani, Hidden Markov models in dynamic system modelling and diagnosis, In : Hidden Markov Models, Theory and Applications, Book edited by Dr. Przemyslaw Dymarski ISBN 978-953-307-993-6, pp. 25-66, April 2011.

[14] http://www.nquran.com/Quran-flex/index.phpm , 23/8/2012 .

[15] Rabiner, L. and Juang, B. -H., Fundamentals of Speech Recognition, PTR Prentice Hall, San Francisco, NJ (1993).

Short Bio Data for the Author

Talaat Wahbi recevied his PhD degree in 2012. He became an assistant professor in the department of computer science, sudan university of scine and technology (SUST). Now he is the PhD program coordinator in the university. His research interest include: Pattern recognition, Image processing, Cryptography, and distributed systems.

 **Ahwag Mohammed Salih Mohammed Jadeed** rececived her BSc degree in computer since from sudan university of scine and technology (SUST), department of IT, also received MSc from same university. She is joind ALT groub in SUST and develop many datasets, also she is one of the developer of Arabic soundex function. Now, she is a lecture in SUST. Her research interest include: A.I and Algorithms.